Parameter Estimation

$a \mapsto P_s$ e.g. $\sum_i a_i \partial_t^i s = \xi$, $P(\xi) = \mathcal{G}(\xi, 1\!\!1)$

$$P_s = \frac{P_\xi(\omega)}{\left|\sum_i a_i (i\omega)^i\right|^2}$$

$$\begin{tikzcd}
a \arrow[d] & \xi \arrow[d] & \\
P_s \arrow[r] & s \arrow[r] & d
\end{tikzcd}$$

$$P(a|d) = \int \mathcal{D}s\, P(a,s|d)$$

$$P(a,s|d) = \frac{P(d|s,a)\,P(s|a)\,P(a)}{P(d)} \propto P(d|s)\,P(s|a)\,P(a)$$

assume $P(a) = \text{const}$, $d = (r * s) + n$, $P(n|s,a) = \mathcal{G}(s,N)$

$$H(d,s,a) \,\widehat{=}\, H(d|s) + H(s|a) + \dots$$

$$= \frac{1}{2}(d-Rs)^\dagger N^{-1}(d-Rs) + \frac{1}{2}\ln|2\pi N|$$

$$+ \frac{1}{2} s^\dagger S^{-1} s + \frac{1}{2}\ln|2\pi S|$$

$$\widehat{=}^{(s,a)} \frac{1}{2}\Big[s^\dagger \underbrace{(S^{-1} + R^\dagger N^{-1} R)}_{D} s - \underbrace{d^\dagger N^{-1} R}_{j^\dagger D D^{-1}} s - s^\dagger \underbrace{R^\dagger N^{-1} d}_{D^{-1} D j} + d^\dagger N^{-1} d + \mathrm{Tr}\ln(2\pi N) + \mathrm{Tr}\ln(2\pi S)\Big]$$

$$\boxed{\begin{aligned} &\ln|A| = \ln\prod_i \lambda_i = \sum_i \ln\lambda_i \\ &A^\dagger = A > 0 \quad = \mathrm{Tr}\ln A \\ &\text{as } A = O\Lambda O^\dagger \quad \mathrm{Tr}A = \mathrm{Tr}\,O\Lambda O^\dagger \\ &\Lambda = \hat{\lambda},\ O^\dagger = O^{-1} \quad = \mathrm{Tr}\,\Lambda O^\dagger O \\ &\qquad = \mathrm{Tr}\,\Lambda \end{aligned}}$$

$$= \frac{1}{2}\Big[(s - \underbrace{Dj}_{m})^\dagger D^{-1}(s - Dj) - j^\dagger D j + d^\dagger N^{-1} d + \mathrm{Tr}\ln(2\pi N) + \mathrm{Tr}\ln(2\pi S)\Big]$$

$$H(d,a) = -\ln\int\mathcal{D}s\, e^{-H(d,s,a)} = \frac{1}{2}\Big[\mathrm{Tr}\ln(2\pi N) + d^\dagger N^{-1} d - \frac{1}{2} j^\dagger D j + \mathrm{Tr}\ln(2\pi S)\Big]$$

$$-\ln \underbrace{\int\mathcal{D}s\, e^{-\frac{1}{2}(s-m)^\dagger D^{-1}\underbrace{(s-m)}_{s'}}}_{\sqrt{|2\pi D|}}$$

$$= \frac{1}{2}\Big[\ln\Big(\frac{|2\pi N|\,|2\pi S|}{|2\pi D|}\Big) + d^\dagger N^{-1} d - \frac{1}{2} j^\dagger D j\Big]$$

$$H(d,a) = H_0 + \frac{1}{2}\left[\ln|D^{-1}S| - d^\dagger N^{-1} R D R^\dagger N^{-1} d\right]$$

$\uparrow$

$\frac{1}{2}\ln|2\pi N| \cdot \frac{1}{2} d^\dagger N^{-1} d$

$$D^{-1}S = (1 + \underbrace{S R^\dagger N^{-1} R}_{Q}) = [1 + Q]$$

$$= H_0 + \frac{1}{2}\int\frac{d\omega}{2\pi}\left[\ln(1 + P_Q(\omega)) - P_d P_N^{-2} P_R P_S (1+P_Q)^{-1}\right]$$

$$= H_0 + \frac{1}{2}\int\frac{d\omega}{2\pi}\left[\ln(1+P_Q) - \frac{P_d}{P_N}\frac{P_Q}{1+P_Q}\right]$$

$P_Q = (P_S P_R P_N^{-1})(\omega) \qquad P_S = \frac{P_Q P_N}{P_R}$

Maximum Likelihood Estimate

$$0 \overset{!}{=} \frac{\partial H(d,a)}{\partial a} = 0 + \boxed{\frac{1}{2}\int\frac{d\omega}{2\pi}\left[\frac{1}{1+P_Q} - \frac{P_d}{P_N}\frac{1+\cancel{P_Q}-\cancel{P_Q}}{(1+P_Q)^2}\right]\frac{\partial P_Q}{\partial a} = 0}$$

assume $a = (P_Q(\omega))_\omega \Rightarrow \frac{\partial P_Q(\omega)}{\partial P_Q(\omega')} = \delta(\omega - \omega')$

$$= \frac{1}{2}\left[\frac{1}{1+P_Q} - \frac{P_d}{P_N}\frac{1}{(1+P_Q)^2}\right]$$

$$\Rightarrow 1 + P_Q = \frac{P_d}{P_N}$$

$$\boxed{P_Q = \frac{P_d - P_N}{P_N} \text{ is MLE} \Rightarrow P_S = \frac{P_d - P_N}{P_R}}$$

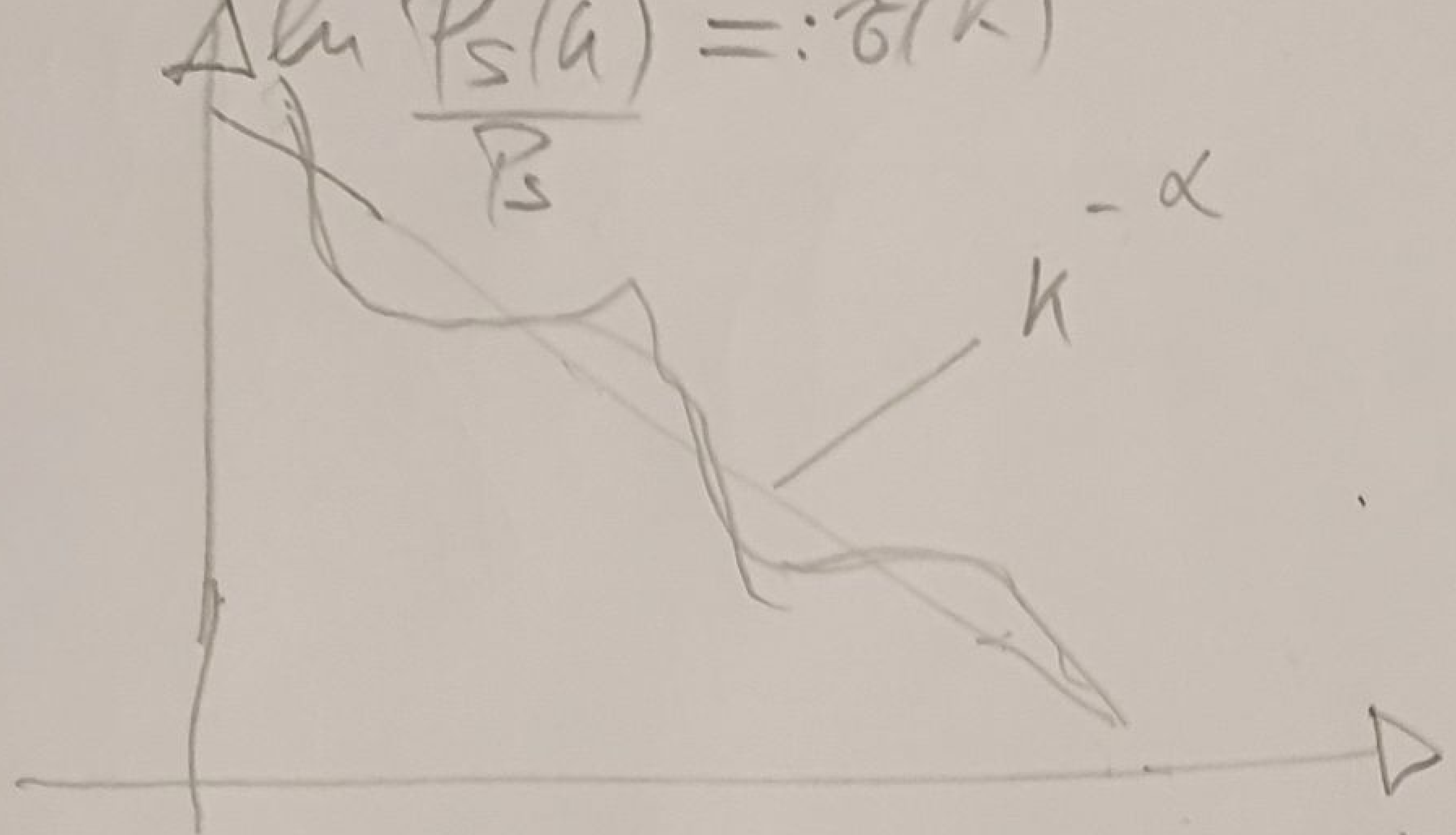
problem: • $P_d - P_N$ can be negative
• overfitting of noise

solution: • use informative prior on P_S, e.g. at $(P_Q(\omega))_\omega$
• use posterior mean instead of noise

• work out $P(s|d) = \int \mathcal{D}a \, \frac{P(d,s,a)}{P(d)}$

Correlated field model

$\ln \frac{P_s(k)}{P_s} =: \tau(k)$



$\ln |k|/k_0 =: x$

$s' = s_0 + s^x$

zero mode $\hookleftarrow \mathcal{G}(s_0 - \bar{s}_0, \sigma^2_{s_0})$

model $\tau(x)$ as GP: $P_s(k) = e^{\tau(\ln|k|)}$

$$\tau(x) = \tau_0 - \alpha x + WP(x) + IWP(x)$$

$$WP(x) = \sigma_{WP} \int_{x_0}^{x} dx' \, \xi^{x'} \qquad \xi \hookleftarrow \mathcal{G}(\xi, \mathbb{1})$$

$$IWP(x) = \sigma_{IWP} \int_{x_0}^{x} dx'' \int_{x_0}^{x'} dx'' \, \eta^{x''} \qquad \eta \hookleftarrow \mathcal{G}(\eta, \mathbb{1})$$

$\tau_0 \hookleftarrow \mathcal{G}(\tau_0 - \bar{\tau}_{ref}, \sigma^2_{\tau_0})$ via fluctuation params

$\alpha \hookleftarrow \mathcal{G}(\alpha - \bar{\alpha}, \sigma^2_{\alpha})$ → Hyperparameters φ

$\sigma_{WP} = e^{\tilde{\sigma}_{WP}}$, $\tilde{\sigma}_{WP} \hookleftarrow \mathcal{G}(\tilde{\sigma}_{WP} - \bar{\tilde{\sigma}}_{WP}, \sigma^2_{\tilde{\sigma}_{WP}})$ asperity

$\sigma_{IWP} = e^{\tilde{\sigma}_{IWP}}$, $\tilde{\sigma}_{IWP} \hookleftarrow \mathcal{G}(\ldots_{IWP}, \ldots)$

resulting model: $\varphi \to (\tau_0, \alpha, \ldots) \to P_s \to s \to d$

highly non-Gaussian and adaptive

→ requires interacting IFT